MATH10282 Introduction to Statistics
Semester 2, 2020/2021
Example Sheet 2

1. A by-election is due to be held next week in a particular town. By polling a sample of the voting population we want to try to predict which of the Labour, Conservative, Liberal Democrat or other candidates will win. Which of the following methods of selection are most likely to yield a representative sample?

  (i) Poll all people of voting age attending a local football game.

  (ii) Poll all the people of voting age who are leaving an expensive town centre restaurant on a particular evening.

  (iii) Obtain a copy of the voter registration list, randomly choose 500 names and question each of them about their voting intentions.

  (iv) Use the results of a telephone poll in which a television station asked its viewers to call and state their voting preference.

  (v) Choose 200 names randomly from the telephone directory and call these people.

2. Which of the following sampling schemes are likely to lead to representative random samples of data? Explain your answer.

  (i) To determine what the average person spends on a vacation, a market researcher interviews passengers aboard a luxury cruise liner.

  (ii) To determine the average income of its graduates 10 years after graduation, a university sends questionnaires in 2015 to all members of the class of 2005 and bases its estimate on the questionnaires returned.

  (iii) To determine public sentiment about certain import restrictions from a particular country, an interviewer asks voters 'Do you feel that this unfair practice should be stopped?'

3. A population consists of the five numbers $2, 3, 6, 8, 11$. Let X_1, X_2 denote respectively the first and second number in a sample of size $n = 2$ chosen by simple random sampling *with replacement*.

  (i) Calculate the population mean.

  (ii) Calculate $\sigma^2 = \text{Var}(X_1)$.

  (iii) By considering all possible samples, calculate the mean and variance of the sampling distribution of $\bar{X}$. Compare your answers with Theorem 1.4 in the lecture notes.

4. With the same population as Question 3 above, by considering all possible samples calculate the mean and variance of the sampling distribution of $\bar{X}$ under sampling *without replacement*. Compare your answer to that for sampling with replacement.

5. This time the population consists of the numbers $3, 7, 9, 11, 15$, which are symmetrically distributed around the value 9. As the distribution here is symmetric, the population mean and median are identical.

    (i) Calculate the population mean.

    (ii) Write down all possible samples of size three which can be drawn randomly *without replacement* from this population. Hence calculate the mean and variance of the sampling distribution of $\bar{X}$.

    (iii) Now use your list of possible samples to calculate the mean and variance of the sampling distribution of the *sample median*.

    (iv) Compare the results obtained in parts (ii) and (iii) above.

6. Suppose that $X_1, \ldots, X_{10}$ is a random sample from a $U[0, 1]$ distribution. Calculate the mean and variance of the sampling distribution of $\bar{X}$.

7. Suppose it is known that a measurement in a population is Poisson distributed with parameter λ, whose value is unknown. However, it is believed that λ is either equal to 10.0 or 12.0. A random sample from this population results in the data

$$(x_1, x_2, x_3, x_4, x_5) = (9, 13, 6, 8, 10, 13) \,.$$

Calculate the joint probability of what has been observed using each of the two plausible values of λ. Which value of λ results in the given sample being most likely to occur?